PTO/SB/21 (04-04) **Application Number** 10/004,131 RANSMITTAL Filing Date October 30, 2001 First Named Inventor Matsunami, Naoto Art Unit (to be used for all correspondence after initial filing) 2182 Examiner Name Nguyen, Mike Attorney Docket Number 16869P-036100US Total Number of Pages in This Submission **ENCLOSURES** (Check all that apply) After Allowance Communication Fee Transmittal Form Drawing(s) to Technology Center (TC) Appeal Communication to Board Fee Attached Licensing-related Papers of Appeals and Interferences Appeal Communication to TC Amendment/Reply Petition (Appeal Notice, Brief, Reply Brief) Petition to Convert to a After Final Proprietary Information Provisional Application Power of Attorney, Revocation Affidavits/declaration(s) Status Letter Change of Correspondence Address Other Enclosure(s) (please Extension of Time Request Terminal Disclaimer identify below): Return Postcard **Express Abandonment Request** Request for Refund Information Disclosure Statement CD, Number of CD(s) The Commissioner is authorized to charge any additional fees to Deposit Certified Copy of Priority Account 20-1430. Document(s) Response to Missing Parts/ Incomplete Application AUG 1 9 2004 Response to Missing Parts Technology Center 2100 under 37 CFR 1.52 or 1.53 SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT Firm Townsend and Townsend and Crew LLP Individual name George B. F. Yee Reg. No. 37,478 Signature Date August 12, 2004 **CERTIFICATE OF TRANSMISSION/MAILING** I hereby certify that this correspondence is being deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on the date shown below. Typed or printed name Cynthia/McKinley Date August 12, 2004

60282954 v1

Signature

日本国特許庁 JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed ith this Office.

出願年月日 Date of Application:

2001年 7月27日

中 爾 釆 号

特願2001-226982

Application Number: [ST. 10/C]:

[JP2001-226982]

制 願 人 wpplicant(s):

株式会社日立製作所

CERTIFIED COPY OF PRIORITY DOCUMENT

特許庁長官 Commissioner, Japan Patent Office 2004年 6月 8日



【書類名】

特許願

【整理番号】

K01000961A

【あて先】

特許庁長官

【国際特許分類】

G06F 12/00

【発明者】

【住所又は居所】

神奈川県川崎市麻生区王禅寺1099番地 株式会社日

立製作所 システム開発研究所内

【氏名】

松並 直人

【発明者】

【住所又は居所】

神奈川県川崎市麻生区王禅寺1099番地 株式会社日

立製作所 システム開発研究所内

【氏名】

八木沢 育哉

【発明者】

【住所又は居所】

神奈川県川崎市麻生区王禅寺1099番地 株式会社日

立製作所 システム開発研究所内

【氏名】

北村 学

【発明者】

【住所又は居所】

神奈川県川崎市麻生区王禅寺1099番地 株式会社日

立製作所 システム開発研究所内

【氏名】

加納 義樹

【発明者】

【住所又は居所】

神奈川県小田原市国府津2880番地 株式会社日立製

作所 ストレージ事業部内

【氏名】

▼高▲本 賢一

【特許出願人】

【識別番号】

000005108

【氏名又は名称】 株式会社日立製作所

【代理人】

【識別番号】

100075096

【弁理士】

【氏名又は名称】

作田 康夫

【手数料の表示】

【予納台帳番号】

013088

【納付金額】

21,000円

【提出物件の目録】

【物件名】

明細書 1

【物件名】

図面 1

【物件名】

要約書 1

【プルーフの要否】

要

【書類名】 明細書

【発明の名称】 記憶装置システム

【特許請求の範囲】

【請求項1】

複数のディスク装置と、

計算機からブロック I / O インタフェースによるアクセスを受けるディスク制御装置と、

計算機からファイル I / O インタフェースによるアクセスを受けるファイルサーバと、

前記ディスク制御装置及び前記ファイルサーバと前記複数のディスク装置との間に接続され、前記ディスク制御装置及び前記ファイルサーバを前記複数のディスク装置に接続するための接続装置とを有することを特徴とする記憶装置システム。

【請求項2】

前記ディスク制御装置及び前記ファイルサーバは、前記複数のディスク装置の うち、自身に割り当てられたディスク装置をアクセスするための情報を保持する 記憶装置を有することを特徴とする請求項1記載の記憶装置システム。

【請求項3】

前記ディスク装置をアクセスするための情報は、前記割り当てられたディスク 装置のアドレス情報を含むことを特徴とする請求項2記載の記憶装置システム。

【請求項4】

前記複数のディスク装置の各々は、自身が割り当てられる前記ディスク制御装置及び前記ファイルサーバを特定するための情報を保持する記憶手段を有することを特徴とする請求項2記載の記憶装置システム。

【請求項5】

前記記憶手段は、前記ディスク装置がデータを記憶するための記憶領域の一部であることを特徴とする請求項4記載の記憶装置システム。

【請求項6】

前記ディスク制御装置及び前記ファイルサーバの各々は、前記複数のディスク

2/

装置の各々に保持された前記特定するための情報を読み出し、自身がアクセス可能なディスク装置を判別することを特徴とする請求項4記載の記憶装置システム。

【請求項7】

前記複数のディスク装置によりストレージプールが構成され、該ストレージプールを管理し、前記ディスク制御装置及び前記ファイルサーバとネットワークを介して接続された管理装置を有することを特徴とする請求項1記載の記憶装置システム。

【請求項8】

前記管理装置は、前記複数のディスク装置ごとに、各々のディスク装置が割り 当てられる前記ディスク制御装置または前記ファイルサーバを対応づけるための ストレージプール管理情報を保持する記憶装置を有することを特徴とする請求項 7記載の記憶装置システム。

【請求項9】

前記複数のディスク装置の各々は、自身が割り当てられる前記ディスク制御装置及び前記ファイルサーバを特定するためのディスク管理情報を保持する記憶手段を有することを特徴とする請求項8記載の記憶装置システム。

【請求項10】

前記管理装置は、システムの初期設定において、前記複数のディスク装置から前記記憶手段に保持された前記ディスク管理情報を取得し、取得した前記ディスク管理情報に基づいて、前記ストレージプール管理情報を生成することを特徴とする請求項9記載の記憶装置システム。

【請求項11】

前記管理装置は、ユーザにより操作される端末装置と接続されており、該端末 装置を介して行われるユーザからの指示に応じて、前記複数のディスク装置の各 々に保持される前記ディスク管理情報の変更、設定を行うことを特徴とする請求 項9記載の記憶装置システム。

【請求項12】

前記ディスク制御装置及び前記ファイルサーバは、前記複数のディスク装置の

うち、自身に割り当てられたディスク装置をアクセスするための使用ディスク情報を保持する記憶装置を有することを特徴とする請求項9記載の記憶装置システム。

【請求項13】

前記使用ディスク情報は、前記割り当てられたディスク装置のアドレス情報を 含むことを特徴とする請求項12記載の記憶装置システム。

【請求項14】

前記管理装置は、前記ストレージプール管理情報に基づいて前記ディスク制御 装置及び前記ファイルサーバの各々に、それぞれが使用可能なディスク装置に関 する情報を通知し、前記ディスク装置及び前記ファイルサーバは、前記通知に基 づいて前記使用ディスク情報を保持することを特徴とする請求項12記載の記憶 装置システム。

【請求項15】

ブロック I / O インタフェースを備えた第1の計算機と、ファイル I / O インタフェースを備えた第2の計算機と、前記第1及び第2計算機と接続され前記第1及び第2の計算機により利用されるデータを格納するための複数のディスク装置を備える記憶装置システムとを有する計算機システムにおいて、前記記憶装置システムは、ストレージエリアネットワークを介して前記第1の計算機と接続されるディスク制御装置と、LANを介して前記第2の計算機と接続されるファイルサーバと、前記ディスク制御装置及び前記ファイルサーバを前記複数の記憶装置と接続するための接続装置と、前記複数のディスク装置の前記ディスク制御装置及び前記ファイルサーバへの割り当てを管理するための管理装置とを備えることを特徴とする計算機システム。

【請求項16】

複数のディスク装置と、

前記複数のディスク装置に格納されたデータに対する計算機からのアクセス要求を処理し、前記複数のディスク装置のうち、自身に割り当てられたディスク装置をアクセスするための情報を保持する記憶装置を備えたコントローラと有することを特徴とする記憶装置システム。

【請求項17】

前記コントローラとして、ファイバチャネルを介して計算機と接続するためのファイバチャネルコントローラを備えるディスクコントローラと、LANを介して計算機と接続するためのネットワークコントローラを有するファイルサーバとを有することを特徴とする請求項16記載の記憶装置システム。

【発明の詳細な説明】

 $[0\ 0\ 0\ 1]$

【発明の属する技術分野】

本発明は、計算機システムで用いられる記憶装置システムに係り、特に、複数種のI/Oインタフェースを有する記憶装置システムに関する。

 $[0\ 0\ 0\ 2]$

【従来の技術】

計算機システムにおいて記憶装置と計算機とを接続するためのインタフェース (以下、I/Fと記す)には、大きく分けて二つのタイプが存在する。第一のタイプは、記憶装置におけるデータの管理単位であるブロックを単位として入出力 (以下、I/Oと記す)のアクセスが行われる「ブロックI/Oインタフェース」である。ブロックI/Oインタフェースには、ファイバチャネル(Fibre Channel)、SCSI(Small Computer Systems Interface)などのインタフェースが含まれる。第二のタイプは、ファイル単位にI/Oのアクセスが行われる「ファイルI/Oインタフェース」である。ファイルI/Oインタフェースは、例えば、ネットワークを介して接続された計算機(ファイルサーバ)により管理される記憶装置のデータを利用する、Network File System(NFS)やCommon Internet File System(CIFS)などのネットワークファイルシステムで用いられる。

[0003]

一方、ネットワークを介して複数の記憶装置を相互に接続する技術も、上述したような I / F の違いによって、2 つのタイプに分類される。一つは、ファイバチャネルを用いて計算機と記憶装置とを接続し、記憶装置への接続のためのネットワークが構築されたストレージエリアネットワーク (SAN) である。第二は

、ローカルエリアネットワーク(LAN)などのネットワークに記憶装置を接続し、ファイル I/O インタフェースで記憶装置をアクセスするネットワークアタッチドストレージ(NAS)である。

[0004]

SANでは、計算機間でのメッセージ交換が行われるネットワークとは別に設けられたストレージ専用の高速ネットワークが用いられる。このため、記憶装置との接続にLANが用いられるNASに比べ、より高速なデータの授受が可能である。SANでは、また、インタフェースとしてブロックI/Oインタフェースが用いられるため、プロトコルオーバヘッドが小さく、高速なレスポンスを得ることができる。この反面、専用のネットワークが必要とされるので、導入の際のコストが大きい。したがって、SANは、主に企業の基幹システムで使用されており、データベースの構築を目的として利用されることが多い。

[0005]

NASでは、LANに記憶装置システムが直接接続される。NASは、既存のLANをそのまま使用できるので、導入の際のコストが小さく、導入が容易である。インタフェースとしては、標準化されたNFS、CIFS等のネットワークファイルシステムを用いられ、ファイル単位でデータを管理できる。このため、データの管理が容易で、かつ複数の計算機で容易にファイルを共有することができる。この反面、NASでは、計算機間の通信が行われるLANを介して記憶装置のアクセスが行われるため、LANの負荷を増大させるおそれがある。また、ネットワークファイルシステムの処理のオーバヘッドが大きく、レスポンスの性能はブロックI/Oより低速となる。このようなことから、NASは、企業のファイル管理システムに主に使用されており、Webコンテンツの管理や、CADのデータファイル管理等のアプリケーションに利用されることが多い。

$[0\ 0\ 0\ 6]$

上述のように、NASとSANはお互いに補完しあう関係にあり、また適用分野も異なる。したがって、一般に、両方は、適所に使い分けられている。

[0007]

SANとNASとの統合に関する技術として、従来、以下に述べる技術が知ら

れている。

[0008]

計算機は、LANにより接続されたファイルサーバを介して記憶装置システムと接続され、あるいは、ファイルサーバと共に、SANを介して記憶装置システムと直接的に接続される。ファイルサーバは、ファイバチャネルなどのインタフェースにより記憶装置システムと接続される。ここで使用される記憶装置システムとしては、複数のディスク装置と、これら複数の記憶装置と計算機との間に介在し、ディスク装置へのデータのアクセスを制御するディスク制御装置とを有するいわゆるディスクアレイシステムがある。

[0009]

記憶装置システムに保持されたデータをアクセスする際、ファイルサーバとのみ接続された計算機は、NFS、CIFSなどのプロトコルを用い、ファイルサーバを介して記憶装置システムのファイルをアクセスする。一方、ファイルサーバとSANの双方に接続された計算機は、記憶装置システムをアクセスする際、ファイルサーバに対してリクエストを発行する。しかし、記憶装置システムから計算機へのデータの転送は、SANを介して直接的に行われる。

$[0\ 0\ 1\ 0]$

この従来技術では、ファイバチャネルにより接続されたファイルサーバと記憶 装置システムとでNASが実現される。つまり、この技術では、SANの上位シ ステムとしてNASが構成される。

$[0\ 0\ 1\ 1]$

【発明が解決しようとする課題】

上述した従来技術によれば、SANとNASとを併用することができる。しかし、NASが記憶装置システムの上位装置となっているので、NASにおけるファイルサーバの処理のオーバヘッドに、記憶装置システムの処理のオーバヘッドが重畳される。したがって、NAS単体で用いるよりもオーバヘッドが大きくなってしまう。

[0012]

また、ファイルの論理的な格納位置は、ファイルサーバにより決定されるが、

7/

記憶装置システム内でのファイルの物理的な格納位置は、記憶装置システムにより決定される。したがって、ファイルサーバは、記憶装置システム内でのファイルの物理的な格納位置を最適化するようなデータ配置することができず、システムの性能が制限されてしまう。

[0013]

さらに、上述した従来技術によれば、NASのみを利用する場合であっても記憶装置システムの他、ファイルサーバとなる計算機が必要であり、初期投資が大きくなってしまう。システムの拡張に関しても、用いられる記憶装置システムの構成により制限されてしまうという問題がある。

[0014]

本発明の目的は、NASとSANを混在させることができ、システム構成に自由度を有する記憶装置システムを提供することである。

[0015]

【課題を解決するための手段】

上記目的を達成するために、本発明の記憶装置システムは、複数のディスク装置と、計算機からブロック I / O インタフェースによるアクセスを受けるディスク制御装置と、計算機からファイル I / O インタフェースによるアクセスを受けるファイルサーバと、ディスク制御装置及びファイルサーバと複数のディスク装置との間に接続され、ディスク制御装置及びファイルサーバと複数のディスク装置とを接続するための接続装置とを有する。本発明の好ましい一つの態様において、ディスク制御装置及びファイルサーバは、複数のディスク装置のうち、自身に割り当てられたディスク装置をアクセスするための情報を保持する。

[0016]

本発明の他の観点によれば、本発明の記憶装置システムは、計算機とファイバチャネルを介して接続するためのファイバチャネルコントローラを備えたディスク制御装置と、計算機とLANを介して接続するためのネットワークコントローラを備えたファイルサーバと、これらディスク制御装置及びファイルサーバにより使用され、計算機が利用するデータを格納する複数のディスク装置とを有する。記憶装置システムは、好ましくは、複数のディスク装置のディスク制御装置及

びファイルサーバへの割り当てを管理する管理装置を含む。管理装置は、ディスク装置の各々について、ディスク装置をそのディスク装置が割り当てられたディスク制御装置あるいはファイルサーバと対応づける情報を保持している。管理装置は、この情報を各ディスク制御装置及びファイルサーバに通知する。各ディスク制御装置及びファイルサーバは、通知された情報に基づき自身が使用するディスク装置を認識し、これに従ってディスク装置を使用する。

[0017]

【発明の実施の形態】

図1は、本発明の一実施形態における計算機システムの構成を示す簡略なブロック図である。

[0018]

本実施形態において、記憶装置システム1(以下、ストレージと呼ぶ)は、ファイバチャネルにより構成されたストレージエリアネットワーク(SAN)8を介して複数のSANクライアント100に、また、ローカルエリアネットワーク(LAN)9を介して複数のNASクライアント110に接続される。ここで、SANクライアント100は、パーソナルコンピュータ、あるいは一般にサーバと呼ばれるような計算機であり、ブロックI/Oインタフェースでストレージ1をアクセスする。一方、NASクライアント110は、SANクライアント100と同様に、パーソナルコンピュータ、あるいは、サーバなどの計算機であるが、ファイルI/Oインタフェースでストレージ1をアクセスする。

[0019]

ストレージ1は、また、管理用LAN13を介して管理端末12に接続される。管理端末12は、ユーザがストレージ1を監視し、あるいは、ストレージ1に対する各種の設定を行うために使用される。

[0020]

ストレージ1は、複数のディスクアレイコントローラ20、複数のファイルサーバ30、複数のディスク装置41、ディスクプール管理装置5、ディスクプール接続装置61を含んで構成される。ディスクプール管理装置5は、複数のディスク装置41の集合により構成される記憶領域を、1つの大きな記憶領域である

ディスクプール4として管理する。ストレージプール4を構成するディスク装置41は、JBOD (Just Bunch Of Disks) とも称される。

[0021]

ディスクアレイコントローラ 2 0、ファイルサーバ 3 0、及びディスクプール管理装置 5 は、管理ネットワーク 7 を介して接続され、相互に通信を行うことができる。また、ディスクアレイコントローラ 2 0 とファイルサーバ 3 0 は、ディスクプール接続装置 6 1 を介してディスクプール 4 を構成するディスク装置 4 1 に接続される。本実施形態では、ディスクアレイコントローラ 2 0 とディスクプール接続装置 6 1 との間は、ファイバチャネルにより接続される。ファイバチャネル接続装置 6 1、ディスクプール管理装置 5、及びディスクプール 4 を構成する核ディスク装置 4 1 との間は、ファイバチャネルによるディスクプール 4 を構成する核ディスク装置 4 1 との間は、ファイバチャネルによるディスクプール接続ネットワークを介して接続される。

[0022]

図2は、ストレージ1の装置構成を示す簡略化された構造図である。

[0023]

ストレージ1は、ラック15にディスクアレイコントローラ20、ファイルサーバ30、ディスクプール管理装置5、ディスクプール接続装置61、及びJBOD筐体140の各コンポーネントを収納して構成される。ラック15は、例えば、19インチの横幅を有しており、各コンポーネントを収納するためのスロットを有する。各スロットは、コンポーネントの厚さの単位の整数倍に相当する高さを有することが望ましい。JBOD筐体140の各々は、複数のスロットを有しており、ディスクプール4を構成するディスク装置41を複数収納することができる。なお、すべてのコンポーネントを1台のラック15に収納する必要はなく、複数のラック15にコンポーネントを分散させて収納してもかまわない。

[0024]

図3は、ディスクアレイコントローラ20の構造を示す模式図である。

[0025]

ディスクアレイコントローラ20は、ディスクアレイコントローラ全体の制御

を行うプロセッサ201、プロセッサ201により実行される制御プログラム、制御データなどが格納されるメモリ202、ディスクアレイコントローラ20をSAN8に接続するためのファイバチャネルコントローラ203、ディスクプール4に接続するためのファイバチャネルコントローラ204、ファイバチャネルコントローラ203と204との間でデータの転送を制御するデータ転送制御コントローラ205、管理ネットワーク7に接続するための管理ネットワークコントローラ206、及び、計算機とディスク装置との間のデータ転送速度の調整、及びデータをキャッシュするために用いられるデータバッファ207を有する。

[0026]

メモリ202には、制御プログラムとして、ディスクアレイ制御プログラム2020、ディスクアレイコントローラ20が使用するディスク装置を特定するためのディスク情報をディスクプール管理装置5から取得するディスク情報取得プログラム2021が保持される。また、メモリ202には、取得したディスク情報が格納されるディスク情報管理テーブル2022が格納されている。

[0027]

図4は、ファイルサーバ30の構成を示す簡略なブロック図である。

[0028]

ファイルサーバ30は、ファイルサーバ30全体の制御を行うプロセッサ301、プロセッサ301で実行される制御プログラム及び制御データが格納されるメモリ302、LAN9に接続するための通信コントローラ303、ディスクプール4に接続するためのファイバチャネルコントローラ304、通信コントローラ303とファイバチャネルコントローラ304の間でデータ転送を制御するデータ転送コントローラ305、及び、管理ネットワーク7に接続するための管理ネットワークコントローラ306を備える。

[0029]

メモリ302には、制御プログラムとして、ファイルサーバ制御プログラム3020、ディスクプール管理装置5からファイルサーバ30が使用するディスク装置41を特定するための情報を取得するディスク情報取得プログラム3021を保持している。メモリ302はまた、取得した使用ディスク装置の情報が格納

されるディスク情報管理テーブル3022を保持する。さらに、メモリ302は、計算機とディスク装置41との間のデータ転送速度の調整、及びファイルデータをキャッシュするためのデータバッファ3023を有する。

[0030]

図5は、ディスクプール4及びディスクプール接続装置61の構成を示す簡略なブロック図である。

[0031]

ディスクプール4は、複数のJBOD筐体140に収納された複数のディスク装置41を有して構成される。ディスク装置41の数は任意であり、必要に応じて、増加、減少させることができる。また、図では、JBOD筐体140が4つ示されているが、この数も任意とすることができる。ディスク装置41は、JBOD筐体140内でループ状に接続される。

[0032]

ディスクプール接続装置 6 1 は、複数の接続ポート 6 1 0、及び複数の拡張ポート 6 1 1 を有する。各接続ポート 6 1 0 には、ディスクアレイコントローラ 2 0、ファイルサーバ 3 0、及び J B O D 筐体 1 4 0 を任意の台数接続することができる。本実施形態では、必要に応じて、拡張ポート 6 1 1 に他のディスクプール接続装置を接続することで、複数台のディスクプール装置 6 1 をカスケード形態で接続することができる。

[0033]

ディスクプール接続装置 6 1 を複数台カスケードに接続することで、自由にポート数を増やすことができる。これにより、ディスクアレイコントローラ 2 0 、ファイルサーバ 3 0 、及び J B O D 筐体 1 4 0 の接続台数や接続構成を自由に変更することが可能となる。したがって、システム構成のスケーラビリティや、システム構成の自由度が著しく向上する。

[0034]

ディスクプール4とディスクプール接続装置61との接続インタフェースにファイバチャネルを適用した場合、ディスクプール接続装置61として、任意の2つの接続ポート間を接続したスターネット構成を有するスイッチ、あるいは、全

ての接続ポートがループ状に接続されたハブの双方を用いることができる。ユーザは、ハブ及びスイッチを、必要とされるデータ転送帯域やコストに合わせて選択することができる。本実施形態では、ディスクプール接続装置 6 1 がハブであるかスイッチであるかを特定せずに説明する。

[0035]

図6は、ディスクプール管理装置5の構成を示す簡略なブロック図である。

[0036]

ディスクプール管理装置 5 は、装置全体の制御を行うプロセッサ 5 1、プロセッサ 5 が実行する制御プログラム、及びプロセッサ 5 1により利用される制御データを格納するメモリ 5 2、管理ネットワーク 7 と接続するための管理ネットワークコントローラ 5 4、ファイバチャネルによりディスクプール 4 と接続するためのファイバチャネルコントローラ 5 4、及び、管理 LAN 1 3 に接続するための通信コントローラ 5 5 を有している。

[0037]

メモリ52は、制御プログラムとして、ディスクプール4を管理するためのディスクプール管理プログラム521、ディスクアレイコントローラ20並びにファイルサーバ30(以下、これらを総じて単にコントローラとも言う)を認識し、それを認証するためのコントローラ認識・認証管理プログラム523、及びストレージ1の管理を行うためのストレージ管理プログラム525を格納している。ディスクプール管理プログラム521は、具体的には、ディスクプール4を構成するディスク装置41を検出し、コントローラに割り当てたディスク装置41の排他制御を行う。

[0038]

メモリ52には、制御データとして、各コントローラへのディスク装置41の 割り当て状況を管理するための情報を保持するディスクプール管理テーブル52 2、コントローラ認識・認証管理プログラム523により検出されコントローラ の管理情報を保持するコントローラ管理テーブル524が格納される。

[0039]

管理LAN13を介してディスクプール管理装置5に接続される管理端末12

は、ユーザに対し各種情報を提示するためのディスプレイ、及びユーザからの情報の入力を受け付けるためのキーボード、マウスなどの入出力装置を備える。管理端末12上では、ブラウザプログラムが動作している。ブラウザプログラムを介して、ディスクプール管理装置5上で動作するストレージ管理プログラム525によりユーザに提供される各種情報の表示、及びストレージ管理プログラム525への各種情報の入力が行われる。

[0040]

図7は、ディスクプール管理プログラム521により実現されるディスクプールの構成認識処理の流れを示すフローチャートである。ディスクプール管理プログラム521は、ユーザによりストレージ1の電源が投入され、あるいは、ストレージ1がリセットされると、プロセッサ51によりその実行が開始される。

[0041]

プロセッサ51は、ディスクプール管理プログラム521の実行を開始すると、全てのディスク装置41がレディ状態になるまで、適当な時間ウェイトする(ステップ6000)。

[0042]

適当な時間ウェイトした後、プロセッサ51は、検索する「パスID」を示す変数 n に値 "0"をセットして初期化する。パスIDとは、ディスク41が接続される1つのループ又はJBOD筐体140を識別するための識別子である。ここでは、パスIDとして、ループ又はJBOD筐体140に割り当てられた番号、もしくはスイッチに接続されたループを代表するWWN(World Wide Name)を用いるものとする(ステップ6002)。続いてプロセッサ51は、パス上の「アドレス」を示す変数 k に値 "0"をセットして初期化する。パス上のアドレスは、各ループ、あるいはJBOD筐体140内で、個々のディスク装置41を特定するため、各ディスク装置41に割り当てられる番号である。パスがループの場合、アドレスは、ループIDに相当する(ステップ6004)。

[0043]

次に、プロセッサ51は、パスID=n、アドレス=kの位置にディスク装置 41があると想定してディスクプール4にInquiryコマンドを発行する(ステッ



プ6006)。プロセッサ51は、発行したInquiryコマンドに対してディスクプール4からの応答の有無を判別する。ディスクプール4からの応答がなければ、プロセッサ51は、パスID=n、アドレス=kの位置に該当するディスク装置が存在しないと判断し、ステップ6016の処理に移り、次のディスク装置41を検索する(ステップ6008)。

[0044]

ディスクプール4からの応答があった場合、プロセッサ51は、パスID=n、アドレス=kの位置にあるディスク装置41から、その容量等の情報を取得する(ステップ6010)。続いて、プロセッサ51は、ステップ6010で容量等の情報を得たディスク装置41に対してリードコマンドを発行する。ここで発行されるリードコマンドは、ディスク装置41の特定の領域を読み出すためのコマンドである。特定の領域には、このディスク装置41を使用しているコントローラに関する情報(以下、コントローラ情報と呼ぶ)が格納されている。コントローラ情報としては、例えば、システム内でコントローラに割り当てられるコントローラ番号、コントローラの製造時にコントローラに付与されるコントローラ識別番号、コントローラ種などの情報が用いられる。プロセッサ51は、リードコマンドに対する応答として、ディスク装置41を使用しているコントローラ情報を取得することができる。この特定の領域にコントローラ情報が記録されていなかった場合、プロセッサ51は、ディスク装置41が未使用であると判定する(ステップ6012)。

[0045]

プロセッサ51は、ディスク装置の容量、コントローラ情報(未使用である場合、それを示す情報)等を取得すると、それらの情報をパスID、アドレスと共にディスクプール管理テーブル522に登録する(ステップ6014)。

[0046]

ステップ6016で、プロセッサ51は、変数 k に1を加算する。続いて、k の値が1パス内の最大のアドレス値を越えたか調べる。k が、最大アドレス値内であれば、ステップ6006以降の処理を繰り返し、当該パスに接続された他のディスク装置41の検査を行う(ステップ6018)。



[0047]

kの値が1パス内の最大アドレス値を越え、パス内のすべてのアドレスについてディスク装置41の検査を終えると、プロセッサ51は、次のパスについて検査を続けるために、変数nに1を加算する(ステップ6020)。続いて、nの値が接続可能なパスのパスIDの最大値を越えていないか判別する。パスIDの最大値を超えていない場合には、ステップ6004以降の処理が繰り返され、新たなパスIDのパスについてディスク装置41の検査が行われる(ステップ6022)。

[0048]

ステップ6022において、nの値がパスIDの最大値を超えていれば、すべてのパスについて検査を終えたことになり、プロセッサ51は、初期設定処理を終了する。

[0049]

本実施形態では、パスIDとアドレスが一定の範囲の番号であるものとし、変数 n、kを "0"から昇順に加算していき、全てのパスID、及び全てのアドレスを検査している。しかし、ディスク装置 4 1 が存在すると予想される位置をすべて過不足なく検索することができれば、別の方法で検査を実施してもよい。

[0050]

図8は、ディスクプール管理テーブル522の構成を概念的に示すテーブル構成図である。

[0051]

ディスクプール管理テーブル522には、初期設定処理で検出されたディスク 装置ごとに、そのパスID5201、アドレス5202、ディスク容量5203 、使用コントローラ識別番号5204、及びディスク装置の状態5205等の情 報が格納される。

[0052]

使用コントローラ識別番号5204は、パスID5201及びアドレス520 2で特定されるディスク装置41が割り当てられたコントローラの識別番号である。ディスク装置41がいずれのコントローラにも割り当てられていない未使用



の状態にあるときは、ブランク、あるいは未使用を示す情報が設定される。

[0053]

図9は、コントローラ認識・認証プログラム523により実現されるコントローラの認識処理の流れを示すフローチャートである。コントローラの認識処理は、ディスクプールの構成認識処理に続けて行われる。また、コントローラの認識処理は、新たにコントローラが増設されたような場合にも実施されてもよい。各コントローラには、あらかじめ管理のためにアドレスが割り当てられている。このアドレスとしては、たとえば、管理ネットワークが一般的なLANの場合、MACアドレスやIPアドレスを用いることができる。

[0054]

プロセッサ51は、コントローラ認識・認証プログラム523による処理を開始すると、全てのコントローラがレディ状態になるまでの適当な時間ウェイトする(ステップ6100)。適当な時間ウェイトした後、プロセッサ51は、アドレスを指定してコントローラに問い合わせを発行する。この問い合わせは、管理ネットワーク7を介して行われる(ステップ6102)。

[0055]

プロセッサ51は、問い合わせに対する応答の有無を判別する(ステップ6104)。問い合わせに対し、コントローラから応答があると、プロセッサ51は、そのコントローラについて、コントローラ番号、コントローラ種類、及びその状態の各情報を取得する。ここで取得されるコントローラ種類の情報は、そのコントローラがディスクアレイコントローラ20であるかファイルサーバ30であるかを示す情報である。また、状態に関する情報は、そのコントローラが正常状態であるのか、あるいは、障害や異常が発生しているのかといった装置の状態を表す情報である(ステップ6106)。続いて、プロセッサ51は、取得した情報をコントローラ管理テーブル524に登録する(ステップ6108)。

[0056]

最後に、プロセッサ51は、すべてのコントローラの検索を終えたかどうか判別する。この判別は、例えば、予め設定されたアドレスの範囲について、その範囲内のすべてのアドレスを用いて上述した処理を行ったか否かを判断することに



より実現できる。コントローラの検索が未完了であれば、ステップ6102以降の処理を繰り返し、次のコントローラを検索する。一方、すべてのコントローラの検索を完了すると、プロセッサ51は、コントローラの認識処理を終了する(ステップ6110)。

[0057]

ステップ6104においてコントローラから応答がなかった場合、プロセッサ51は、問い合わせを行ったアドレスを持つコントローラが存在しないと判断し、ステップ6106、6108の処理をスキップし、ステップ6110の処理に移る。

[0058]

なお、コントローラのアドレスは、ラック15のスロット位置にコントローラを搭載することによって自動的に付与されるようにすることもできる。また、コントローラ番号は、システム内で唯一である必要がある。具体的には、コントローラの製造番号などのように、その装置に唯一無二な番号をコントローラ番号として用いることが好ましい。

[0059]

図10は、コントローラ管理テーブル524の構成を概念的に示すテーブル構成図である。

[0060]

コントローラ管理テーブル524には、検出されたコントローラの管理ネットワーク7におけるアドレス5241、及びステップ6106で取得されたコントローラ番号5242、コントローラ種類5243、並びに状態5244等の情報が格納される。

$[0\ 0\ 6\ 1]$

ストレージ1の電源の投入を契機に、各コントローラ(ディスクアレイコントローラ20、ファイルサーバ30)は、制御プログラム(ディスクアレイ制御プログラム2020、ファイルサーバ制御プログラム3020)の実行を開始し、自身の初期設定を行う。

[0062]



初期設定を終了した後、各コントローラは、ディスク情報取得プログラム2021、3021を起動し、自身が使用するディスク装置に関する情報の取得を行う。図11は、このとき行われるディスク認識処理の大まかな流れを示すフローチャートである。なお、図11では、コントローラ側での処理のフローを左側に、ディスクプール管理装置側での処理のフローを右側に示している。以下では、ディスクコントローラ20での処理を例に説明するが、ファイルサーバ30においても同様の処理が行われる。

[0063]

プロセッサ201は、自コントローラの識別番号を取得する(ステップ620 0)。プロセッサ201は、取得した識別番号を用いて、ディスクプール管理装置5に対し認証の問い合わせを行う。この問い合わせは、管理ネットワーク7を 経由して実施される(ステップ6202)。プロセッサ201は、問い合わせ後 、ディスクプール管理装置5において認証が完了するまで処理をウェイトさせる (ステップ6204)。

[0064]

問い合わせを受けたディスクプール管理装置 5 のプロセッサ 5 1 は、コントローラ認識・認証プログラム 5 2 0 1 を起動し、認証処理を実行する。認証処理においてプロセッサ 5 1 は、問い合わせのあったディスクアレイコントローラ 2 0 の情報がコントローラ管理テーブル 5 2 4 に登録されているかどうか確認し、その認証を行う(ステップ 6 2 5 1)。プロセッサ 5 1 が他のコントローラの認識処理を実施中である場合、全てのコントローラについての処理が終了するまでこの問い合わせの処理は待たされる。問い合わせのあったディスクアレイコントローラ 2 0 の情報がコントローラ管理テーブル 5 2 4 に登録されていることが確認できると、プロセッサ 5 1 は、認証済みであることを示す情報を問い合わせもとのディスクアレイコントローラ 2 0 に返送する(ステップ 6 2 5 3)。

[0065]

認証済みであることを示す情報を受け取ったディスクアレイコントローラ20 のプロセッサ201は、自コントローラで使用するディスク装置41を特定する ため、自身の識別番号を含む使用ディスクの問い合わせをディスクプール管理装



置5に対して送信する。この問い合わせは、管理ネットワーク7を経由して行われる(ステップ6206)。プロセッサ201は、使用ディスクの問い合わせが 完了するまで処理の実行をウェイトさせる(ステップ6208)。

[0066]

問い合わせを受けたプロセッサ51は、ディスクプール管理プログラム521を実行し、この問い合わせを処理する。具体的にプロセッサ51は、問い合わせに含まれる識別番号をキーとしてディスクプール管理テーブル522を検索し、問い合わせ元のコントローラが使用するディスク装置41を特定する。その後、プロセッサ51は、特定したディスク装置41のパスID、アドレス、ディスク容量、及び状態の各情報を含む一覧リストを作成する(ステップ6255)。一覧リストを作成すると、プロセッサ51は、その一覧リストを問い合わせもとのコントローラに返送する(ステップ6257)。

[0067]

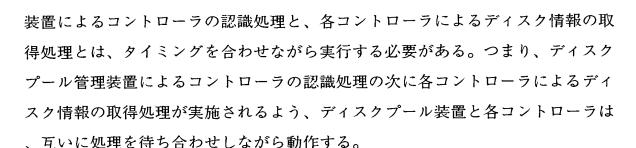
プロセッサ201は、ディスクプール管理装置5から使用ディスクの情報を一覧リストの形態で受信する(ステップ6210)。プロセッサ201は、受信した一覧リストに登録されている情報に基づいて、ディスク情報管理テーブル202に使用するディスク装置41に関する情報を登録し、処理を終了する(ステップ6212)。

[0068]

使用ディスク情報を取得したディスクアレイコントローラ20、及びファイルサーバ30は、ディスク情報管理テーブル2022、3022に登録されたディスク装置41が自コントローラで使用するすべてのディスク装置であると認識して、それらのディスク装置41を使用する。このようにすることで、1つのディスクプール4を構成するディスク装置41を各コントローラが排他的に使用することができる。

[0069]

なお、上述した各処理のうち、ディスクプール管理装置によるディスクプール の構成認識処理と、各コントローラにおける初期化処理は、互いに同時、並行し て実施することができる。一方、これらの処理の後行われるディスクプール管理



[0070]

図12は、ディスクプール4を構成するディスク装置41のコントローラへの 割り当ての一例を示す模式図、図13は、図12に示した割り当ての様子を、コ ントローラ毎に論理的に整理した構成を示した図である。ここで示した例では、 40台のディスク装置41が、それぞれ排他的に各コントローラに割り当てられ ている。

[0071]

図12において、ディスク装置41を表すシンボル中にある "Array#" は、そのディスク装置41が#番目のディスクアレイコントローラ20に割り当てられていることを示している。同様に、 "FSVR#" のディスク装置41は、#番目のファイルサーバ30に割り当てられていることを示している。

[0072]

一方、図13において、ディスク装置41を表すシンボル中の"n-k"は、そのディスク装置が、パス ID "n"に接続されたアドレス"k"のディスク装置であることを表している。

[0073]

新しいディスク装置41をコントローラに割り当てる場合、ユーザは、管理端末12を用い、ディスクプール管理装置5上のストレージ管理プログラム525を動作させる。ストレージ管理プログラム525は、例えば、図13に示すような、ディスクプール4の論理的な構成を管理端末12のディスプレイに表示する。

[0074]

このような画面上でユーザは、マウスを使用し、未使用を示す領域にあるディスク装置を示すアイコンをドラッグする。そして、ディスク装置41を追加した



いコントローラを示す領域にそのアイコンを移動させる。

[0075]

管理端末12でのユーザの操作によりディスク装置41の割り当てについての情報を受け取ったディスクプール管理装置5は、ディスクプール管理プログラム521を実行して以下の処理を行う。

[0076]

すなわち、ディスクプール管理装置 5 は、管理端末 1 2のディスプレイ上で操作の対象とされたアイコンに対応するディスク装置 4 1 の定められた記憶領域に、そのアイコンが移動された領域に対応するコントローラの識別番号を書き込む。さらに、ディスクプール管理装置 5 は、そのディスク装置 4 1 にInquiryコマンドを発行し、応答として取得した情報をディスクプール管理テーブル 5 2 2 に登録する。この後、ディスクプール管理装置 5 は、管理ネットワーク 7 を介しコントローラにディスク装置 4 1 が追加されたことを通知する。

[0077]

ディスクプール管理装置 5 からディスク装置 4 1 の追加を知らせる通知を受信すると、コントローラは、先に説明したディスク情報取得処理を実行し、追加されたディスク装置 4 1 を特定し、それを使用することができるようになる。

[0078]

以上説明したようにストレージを構成することで、1台のストレージに複数のディスクアレイコントローラ及び複数のファイルサーバを混在させ、SANとNASとが混在したシステムを構築することができる。ストレージの管理は、管理端末から集中的に行うことが可能であり、管理の容易性が向上して管理コストを低減することができる。

[0079]

SANやNASを使用する場合、SANに接続された記憶装置を論理的にまとめて1台の大容量の記憶装置として扱うよりも、複数の容量の小さな記憶装置の集合体と考えた方が使いやすいことも多い。たとえば、会社の各部門ごとに、1台のファイルサーバを使用する場合、各計算機(サーバ)毎にディスクアレイを独立に割り当てたい場合などがある。また、インターネットデータセンタ(ID



C) におけるレンタルサーバやホスティングサービスと呼ばれる形態を構築する場合、及びストレージサービスプロバイダ(SSP)と呼ばれるディスク容量をレンタルするサービスのような形態を構築する場合など、様々な場面でユーザ又は計算機毎にストレージシステムを排他的に割り当てて使用したい場合がある。以下に、このように1つのストレージシステムを論理的に分割して、それぞれの部分を排他的にユーザに割り当てる運用形態の一例について説明する。

[0080]

図14は、ストレージプール4を論理的に分割して運用する際の運用形態の一例を示す模式図である。

[0081]

160~163は、ストレージプール4を論理的に分割して得られる1つのディスクアレイをディスクアレイコントローラ20で制御する仮想ディスクアレイ、170~172は、ストレージプール4を論理的に分割して得られる1つのディスクアレイをファイルサーバ30で制御する仮想ファイルサーバである。

[0082]

ディスク装置 4 1 の割り当て、及び排他管理はディスクプール管理装置 5 で行われる。しかし、ひとたびディスク装置 4 1 がコントローラに割り当てられると、コントローラは、従来のディスクアレイ装置及びファイルサーバ装置と同様に、割り当てられたディスク装置 4 1 を占有することができる。

[0083]

上述した実施形態によれば、SANすなわちブロックI/Oインタフェースを備えるディスクアレイ、及びNASすなわちファイルI/Oインタフェースを備えるファイルサーバとの混在を実現する記憶装置システムを構築することが可能になる。

[0084]

ファイルサーバに直接ディスクが接続されるため、ファイルサーバ利用時においても余分なオーバヘッドが発生せず、高速なファイルのアクセスが実現できる。また、ファイルサーバが直接ディスクを操作するので、ファイルの最適配置も可能となり、さらに高速なファイルアクセスを実現することが可能である。



[0085]

また、上述した実施形態によれば、ディスクアレイコントローラ、ファイルサーバ、ディスク装置などを必要に応じ必要な台数づつ追加することが容易にできる。このため、必要な部位を逐次導入していくことが可能となり、低価格なシステムを提供でき、スケーラビリティや構成自由度に優れたストレージシステムを構築できる。

[0086]

さらに、SANとNASとで同一のディスクプールをシェアしながら各々が使用するディスクの排他制御を行うことができ、分散システムでありながら、集中管理を実現できるので、容易な使用を実現するとともに管理コストの低減を実現できる。

[0087]

またさらに、複数台のSAN、NASが各々仮想的に独立しているように見せかけることができるので、ユーザや計算機毎に独立した分散環境を集中管理の下実現できる。

[0088]

上述した実施形態では、ストレージにディスクプール管理装置を設け、各コントローラが利用できるディスク装置の割り当て、排他利用をディスクプール管理 装置で集中管理している。このような集中管理の形態に代えて、各コントローラ でディスクプールの管理を分散して行うことも可能である。

[0089]

ディスクプールの管理を分散して行う場合も第一実施形態と同様、ディスク装置の特定の領域に、そのディスク装置を使用するコントローラの識別番号が格納される。分散管理を行う場合、ディスクプール管理装置が各コントローラへのディスク装置の割り当てに先立ちこれらの情報を事前に読み出すことに代えて、すべてのコントローラがすべてのディスク装置の上述した特定の領域を検索する。各コントローラは、自身の識別番号を検出すると、そのディスク装置を自身が使用可能なディスクとし判断する。

[0090]

新しいディスク装置を搭載する際には、各コントローラに用意したストレージ 管理プログラムで、ユーザの指示の下、この識別情報を格納するようにすればよい。

[0091]

以上のような構成とすることで、ディスクプール管理装置が不要となり、さらなる小型化、低価格化を実現することができる。

[0092]

【発明の効果】

本発明のによれば、NASとSANを混在させたストレージシステムを構築し、SANとNASの統合的な管理を容易に実現することができる。

【図面の簡単な説明】

【図1】

一実施形態における計算機システムの構成を示す簡略なブロック図である。

[図2]

ストレージの装置構造を示す模式図である。

【図3】

ディスクアレイコントローラの構成を示す簡略なブロック図である。

【図4】

ファイルサーバの構成を示す簡略なブロック図である。

【図5】

ディスクプール及びディスクプール接続装置の構成を示す簡略なブロック図である。

【図6】

ディスクプール管理装置の構成を示す簡略なブロック図である。

【図7】

ディスクプール管理プログラムにより実現されるディスクプールの構成認識処理の流れを示すフローチャートである。

【図8】

ディスクプール管理テーブルの構成を概念的に示すテーブル構成図である。



コントローラ認識・認証プログラムにより実現されるコントローラの認識処理 の流れを示すフローチャートである。

【図10】

コントローラ管理テーブルの構成を概念的に示すテーブル構成図である。

【図11】

ディスク認識処理の流れを示すフローチャートである。

【図12】

ディスクプールを構成するディスク装置のコントローラへの割り当ての一例を 示す模式図である。

【図13】

図12に示した割り当ての様子を、コントローラ毎に論理的に整理した構成を示した模式図である。

【図14】

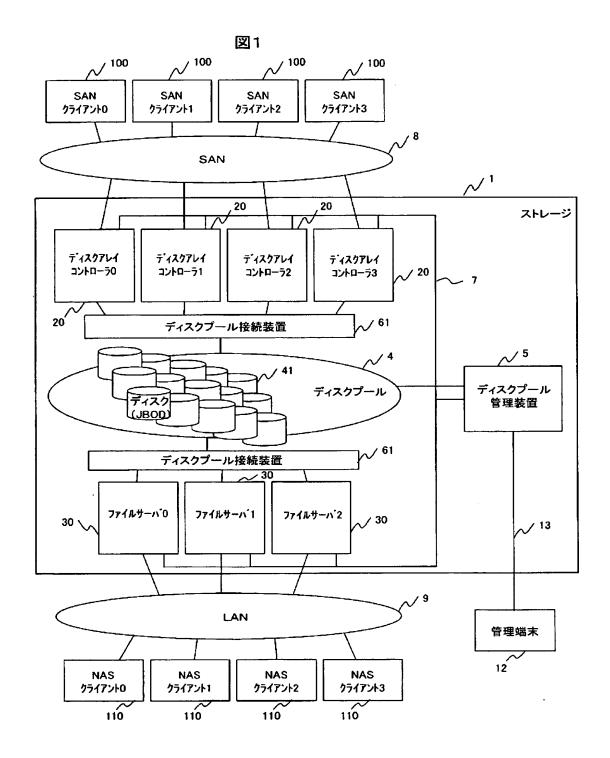
ストレージプールを論理的に分割して運用する際の運用形態の一例を示す模式 図である。

【符号の説明】

1…ストレージ、20…ディスクアレイコントローラ、30…ファイルサーバ、4…ディスクプール、41…ディスク装置、5…ディスクプール管理装置、61…ディスクプール接続装置、7…管理ネットワーク、8…ファイバチャネル、9…イーサネット、100…SANクライアント、110…NASクライアント、12…管理端末、13…管理用LAN、140…JBOD筐体、15…ラック、2020…ディスクアレイ制御プログラム、2021…ディスク情報取得プログラム、2022…ディスク情報管理テーブル、3020…ファイルサーバ制御プログラム、3021…ディスク情報で理テーブル、521…ディスクプール管理プログラム、522…ディスクプール管理テーブル、523…コントローラ認識・認証プログラム、524…コントローラ管理テーブル、525…ストレージ管理プログラム。

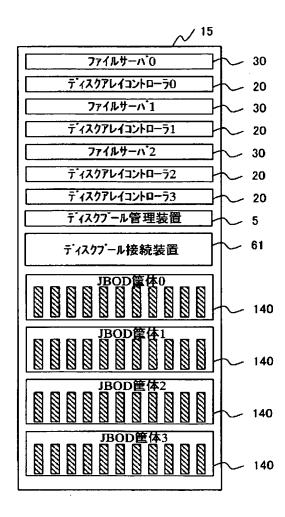


【図1】

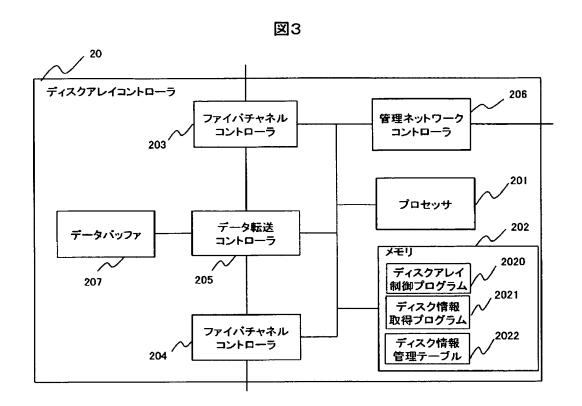


【図2】

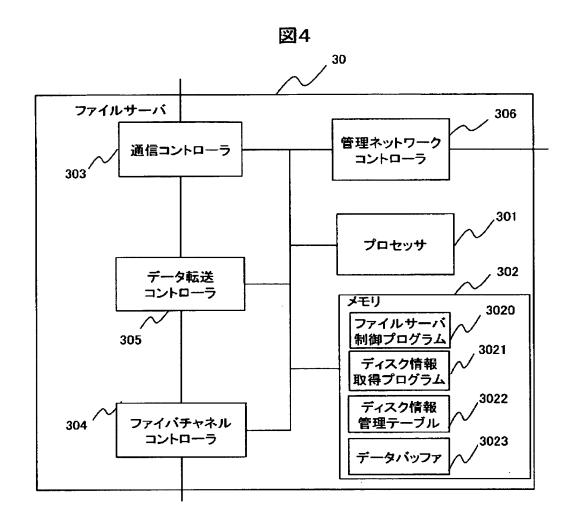
図2



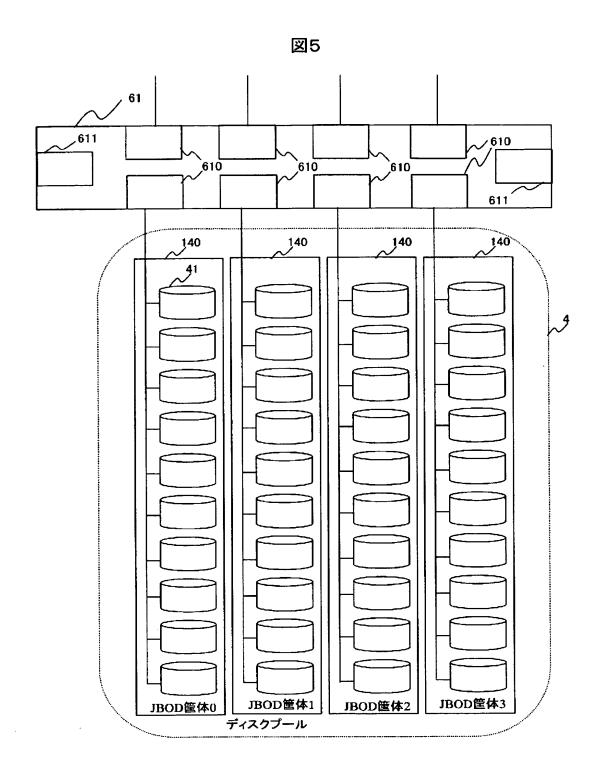
【図3】



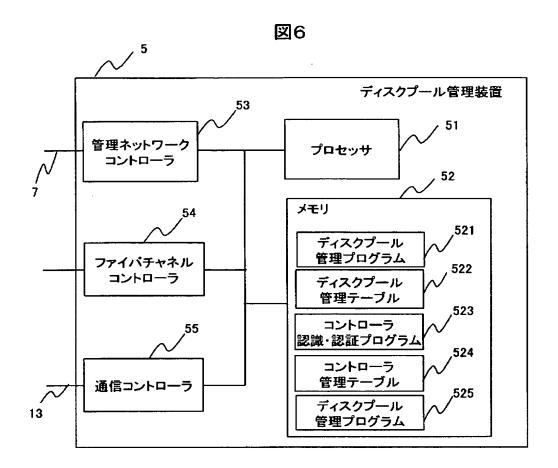
【図4】



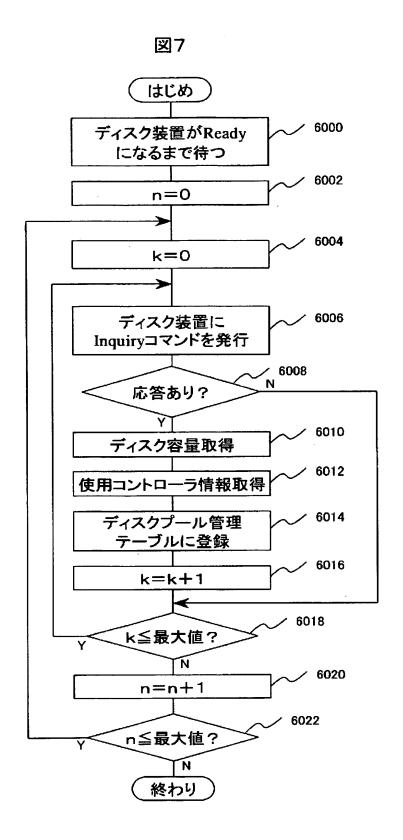
【図5】



【図6】



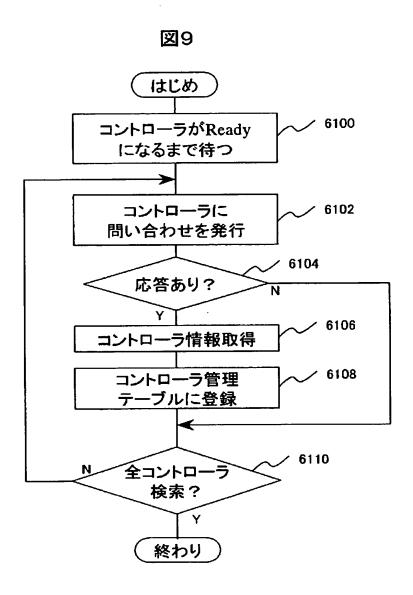
【図7】



【図8】

図8					
~	5201	5202 5203	5204	~	5205
パスID	アドレス	ディスク容量	使用コントローラ識別番号	状態	
0	0	36GB	ARRAY500A1234	正常	
0	1	36GB	ARRAY500A1234	正常	
0	2	36GB	FSVR100B0013	正常	
••••	:		:	:	

【図9】



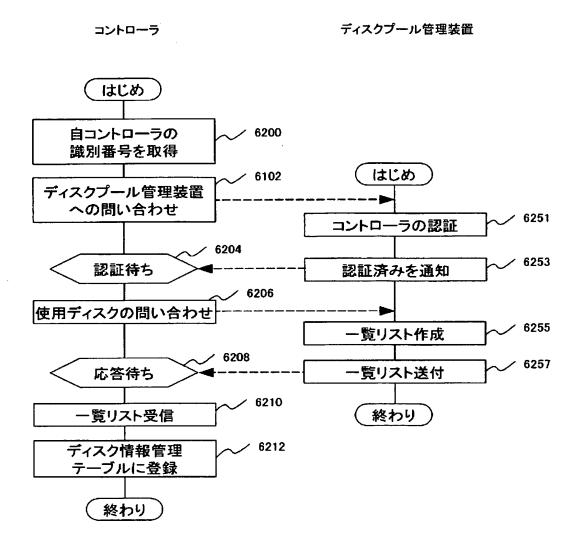
【図10】

図10

524	5243	241 5242	,~/ 5
状態	コントローラ種	使用コントローラ識別番号	アドレス
正常	ファイルサーバ	FSVR100B0013	0
正常	ディスクアレイコントローラ	ARRAY500A1234	1
正常	ファイルサーバ	FSVR100B0059	2
正常	ディスクアレイコントローラ	ARRAY500A2199	3
正常	ファイルサーバ	FSVR100B0153	4
正常	ディスクアレイコントローラ	ARRAY500A1898	5
正常	ディスクアレイコントローラ	ARRAY500A1936	6
	ディスクアレイコントローラ ファイルサーバ ディスクアレイコントローラ	ARRAY500A2199 FSVR100B0153 ARRAY500A1898	3 4 5

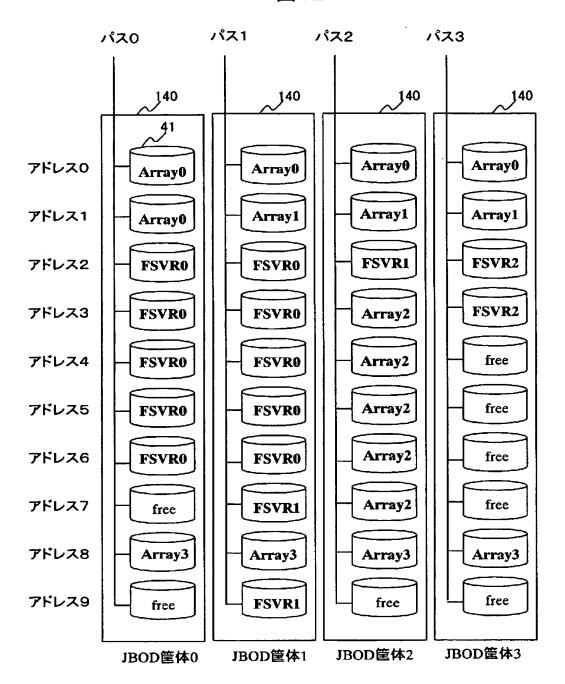
【図11】

図11



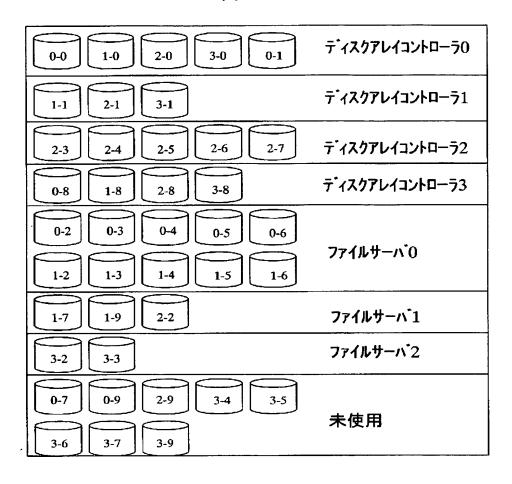
【図12】

図12

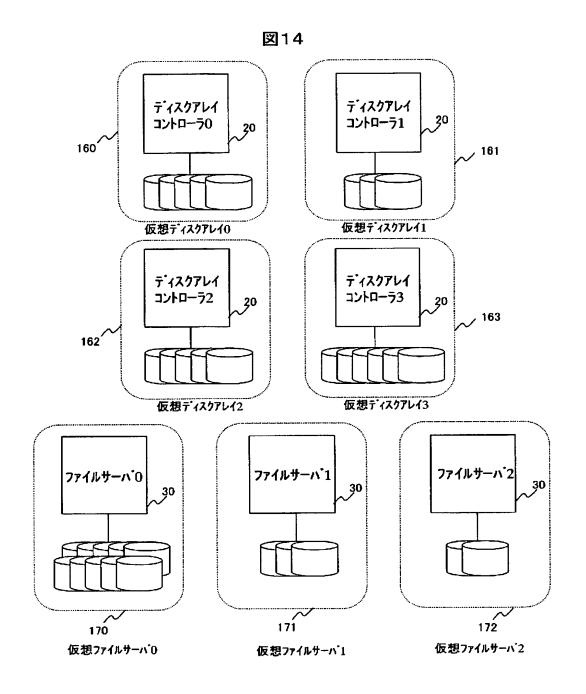


【図13】

図13









【要約】

【課題】

NASとSANを混在させたストレージシステムを構築し、SANとNASの 統合的な管理を容易にする。

【解決手段】

計算機からSAN8を介してアクセスを受け付ける複数のディスクコントローラ20と、計算機からLAN9を介してアクセスを受け付ける複数のファイルサーバ30とを設け、これらディスクコントローラとファイルサーバを、接続装置を介してディスクプール4を構成する複数のディスク装置41と接続する。ディスクプール管理装置5は、ディスクコントローラおよびファイルサーバと、それらの各々が使用するディスク装置との対応関係を管理し、ディスクコントローラおよびファイルサーバに対し、それらに割り当てられたディスク装置の情報を通知する。各ディスクコントローラおよびファイルサーバは、通知された情報に従ってディスク装置を使用する。

【選択図】 図1

認定・付加情報

特許出願の番号 特願2001-226982

受付番号 50101101500

書類名 特許願

担当官 第七担当上席 0096

作成日 平成13年 7月30日

<認定情報・付加情報>

【提出日】 平成13年 7月27日

特願2001-226982

出願人履歴情報

識別番号

[000005108]

1. 変更年月日

1990年 8月31日

[変更理由]

新規登録

住所

東京都千代田区神田駿河台4丁目6番地

氏 名

株式会社日立製作所